Behavioral Experiments in Networked Trade

J. Stephen Judd * Michael Kearns †

ABSTRACT

We report on an extensive series of highly controlled human subject experiments in networked trade. Our point of departure is a simple and well-studied bipartite network exchange model, for which previous work has established a detailed equilibrium theory relating wealth to network topology. A notable feature of this theory is its prediction that there may be significant local variation in equilibrium wealths and prices purely as a result of structural asymmetries in the network.

Our experiments mix recent lines of thought from algorithmic game theory, behavioral economics and social network theory, and are among the first and largest behavioral economics experiments on network effects conducted to date. They continue a line of human subject experiments on networked games and optimization allowing only local interactions.

Categories and Subject Descriptors: J.4 [Social and Behavioral Sciences]: Economics

General Terms: Economics

Keywords: Network Economics, Behavioral Economics

1. INTRODUCTION

In recent years there has been much research on network-based models in game theory, in both the computer science and economics communities. Topics receiving considerable attention include the effects of network topology on equilibrium properties [6, 9, 11, 7, 14], price of anarchy analyses of selfish routing and other networking problems [16], game-theoretic models of network formation (see [17] and citations therein), equilibrium computation in networked settings (see [10, 15] and citations therein), and many others. This large and growing literature has been almost [3, 5] exclusively theoretical, with essentially no accompanying empirical or behavioral studies examining the relevance of the mathematical models to actual behavior.

In this paper, we report on an extensive series of highly controlled human subject experiments in networked trade. Our point of departure is a simple and well-studied bipartite network exchange model, for which previous work has established a detailed equilibrium theory relating wealth to network topology [9]. A notable feature of this theory is its prediction that there may be significant local variation in equilibrium wealths and prices purely as a result of structural asymmetries in the network. A (centralized) efficient algorithm for computing such equilibria in our model was discovered only recently [4], and it uses linear programming as a subroutine. One can view our experiments as a test of human subjects' ability to solve this challenging problem in a distributed setting using only local information.

Our methodology and experiments mix recent lines of thought from algorithmic game theory, behavioral economics and social network theory, and are among the first and largest behavioral experiments on network effects conducted to date. We adopt many of the practices of behavioral game theory [2], which has tended to focus on two-player or small-population games rather than larger networked settings. The work described here continues a broader line of research in behavioral games on networks [12].

In each of our experiments, three dozen human subjects simultaneously engage in trade over an exogenously imposed bipartite network. Trade is permitted only between network neighbors and takes place via a limit order mechanism; communication thus occurs only through local price discovery throughout the network. Subjects have real financial incentives to trade their initial endowments for as much neighboring commodity as possible.

Our results include a detailed examination of how network structure influences collective and individual outcomes (wealths), as well as a study of the behavioral relevance of the equilibrium theory for the model. The networks imposed are drawn from models common in social network theory, including preferential attachment graphs, random (Erdos-Renyi) networks, and some carefully designed structures. Among our most striking findings are the following:

- The behavioral collective performance, as measured by social welfare, is generally strong, with market efficiency close to 90% across all experiments.
- Different network topologies can be strongly distin-

*Department of Computer and Information Science, University of Pennsylvania, Philadelphia PA. Email address: sjudd@seas.upenn.edu

†Department of Computer and Information Science, University of Pennsylvania, Philadelphia PA. Email address: mkearns@cis.upenn.edu

EC'08, July 8–12, 2008, Chicago, Illinois, USA.

guished statistically by individual wealth levels across experiments and other measures.

- When equilibrium theory predicts that greater disparity in individual wealths is required to clear the market, the collective performance (social welfare) of the subjects degrades. However, greater equilibrium wealth disparity does in fact lead to greater behavioral wealth disparity.

- The equilibrium theory appears highly relevant to the behavioral outcomes, outperforming plausible alternatives in predicting individual wealth on an experiment-by-experiment basis. However, our best predictor of behavioral individual wealth nudges the equilibrium theory towards equality (uniform wealth distribution); this can be viewed as a networked form of *inequality aversion*, a bias extensively documented and studied in the behavioral game theory literature on two-player games [2].

We also provide analyses of the behavioral dynamics within the experiments, including the phenomenon of early trading and the fragmentation of liquidity due to trading. All of the above results and almost all others reported are highly statistically significant.

2. THEORETICAL BACKGROUND

The behavioral experiments described in this paper are directly based on a theoretical model for the networked exchange of goods. In this section we describe that model and the detailed equilibrium theory that has been developed for it in previous work; more background on this model and its accompanying theory can be found in [9].

The networked exchange model consists of two equal-sized populations of participants that we shall call the *Milk* players and the *Wheat* players, and a bipartite graph G between them. Milk players begin with 1 fully divisible unit[1] of an abstract commodity that we shall naturally call milk; and Wheat players begin with 1 fully divisible unit of an abstract commodity that we shall naturally call wheat[2]. Mutual interest in trade between Milk and Wheat players is induced via the utility functions: Milk players have *no* utility for milk, and linear utility for wheat; Wheat players have no utility for wheat, and linear utility for milk. Thus each player has a utility function incentivizing them to trade their entire initial endowment for as much of the "other" good as they can.

So far we have described a very special case of classical general equilibrium exchange models such as those of Fisher and Arrow-Debreu. Things get more interesting when we consider the network constraints. The semantics of the bipartite graph G are that players may only exchange goods with their immediate neighbors in the graph, with no resale permitted. To connect our model to the more traditional ones, note that the graph constraints can easily be encoded in the utility functions by having the utilities to (say) a Wheat player of milk bought from its neighbors in G be all equal and linear, and having the utility of milk bought from non-neighbors be zero [8]. This ensures that at equilibrium trade takes place only over the edges of G.

In classical exchange models, all players are free to trade with all other players (which in our setting corresponds to G being the complete bipartite graph between the two populations), which results in all players charging the same (universal) price at equilibrium. The most interesting aspect of our model is that now equilibrium prices may vary across the population due to network asymmetries. To describe equilibria, we need to define:

- A *price*, p_i, for each player i. If i is a Wheat player, p_i is interpreted as the amount of milk that i demands per unit of wheat, with a symmetric interpretation for Milk players.

- A *trading volume*, x_{ij}, for each edge (i, j) in the bipartite graph G, indicating what quantity of the "other" good player i is *receiving* from its neighbor j. Obviously, at equilibrium x_{ij} and x_{ji} must be consistent with the prices p_i and p_j.

With this notation, an equilibrium for our model obeys the following standard properties:

- *Market Clearing.* For every player i, $\sum_{j \in N(i)} x_{ji} = 1$, where $N(i)$ is the set of neighbors of i in G. Thus, the demand from its neighbors exactly exhausts the endowment of player i.

- *Utility Maximization.* For every player i and every $j \in N(i)$, $x_{ij} > 0$ implies that $p_j = \min_{j' \in N(i)}\{p_{j'}\}$. Thus, i receives a non-zero amount of the other good from j only if j is offering the best price available to i (for which there may be ties).

Previous work has established that such an equilibrium always exists for any bipartite graph G [8]. Note that since all endowments are equal to 1 unit, the equilibrium *wealth* (amount of the other good) that any party i receives is exactly equal to their price p_i; more generally, wealths will always be proportional to prices. Furthermore, if $x_{ij} > 0$ for any edge (i, j) in G, it must be the case that $p_i = 1/p_j$. Note that market-clearing implies the average wealth must be 1, or more generally equal to the average initial endowment.

The main purpose of the model above is to examine the simplest setting in which the only asymmetries in wealth arise exclusively due to the network structure, since all utility functions are symmetric and all endowments equal. Indeed, the point of departure for this paper is a series of results carefully quantifying how the equilibrium wealth distribution depends on the topology of G, the most basic being the following:

THEOREM 2.1. *[9] In the bipartite exchange model described above, the maximum wealth at equilibrium is equal to* $\max_S\{|S|/|N(S)|\}$*, where the maximum is taken over all subsets S of players, and $N(S)$ denotes the set of all neighbors of S in G. The minimum wealth will be the reciprocal quantity. In particular, there will be no variation in wealth at equilibrium if and only if G contains a perfect matching subgraph.*

The theorem above can actually be iterated to obtain an inefficient but intuitive algorithm for computing equilibrium wealths. If the vertex set S^* achieves the maximum, the vertices in $N(S^*)$ will enjoy the highest wealth of

[1] For cognitive purposes, in the actual experiments subjects were provided 10.0 fully divisible units each, but it is easier to describe the theory with 1-unit endowments. The difference is entirely cosmetic.

[2] We deliberately avoid the use of a good called "cash" and the terminology "buyers" and "sellers" due to the confusion and unintended interpretations they might introduce.

$|S^*|/|N(S^*)| \geq 1$, while those in S^* will suffer the reciprocal wealth. It can be shown that we can now remove S^* and $N(S^*)$ from G and repeat to find the remaining wealths in a similar fashion. A considerably faster algorithm [4] is used for the actual equilibrium computations discussed later.

More importantly, the theorem above has been applied to a variety of common stochastic models for network formation in order to see the implications for wealth distribution. For instance, it has been shown that bipartite Erdos-Renyi graphs of N vertices yield essentially no variation in price in the limit of large N, whereas models such as preferential attachment yield max/min wealth ratios that are a root of N and thus unbounded as N grows [9]. Such findings are directly relevant to the networks we selected for our experiments, as discussed in the next section.

3. NETWORK TOPOLOGIES

The bipartite graphs chosen for our behavioral experiments were selected in order to provide direct tests of the equilibrium theory just described, and also to sample the diversity of stochastic network formation models that have been proposed in recent years (see for instance [18]). We begin by describing the network topologies, followed by discussion of their equilibrium properties.

The networks used in our 28 experiments were divided into 10 different topological "families", samples of which are given in Figure 1, where color coding is used to distinguish the two populations. The simplest of these, the *pairs* topology, matches every Wheat player with a unique Milk player, resulting in 18 disconnected pairs of players; here there is thus no "network" in the usual sense. The topology *cycle2* is a simple cycle in which Milks and Wheats alternate around the cycle. The topology *cycle4* begins with a simple cycle but then adds edges between all players 3 hops apart, giving every party 4 trading partners rather than 2.

The *clan* topology is also highly structured. It consists of four imbalanced "sub-economies" of 8 players each — 6 of one type, 2 of the other — in which all edges are present. These four sub-economies have no direct connections; rather, they are all fully connected to four favored "super-trader" players, two of each type, who can trade with any party of the opposing type. The *clanp5* topology includes all edges of clan and adds each missing Milk-Wheat edge in the clan topology with probability 0.05, while *clanp10* does so with probability 0.1. Both of these stochastic families were sampled 3 times to yield 3 different *clanp5* graphs and 3 different *clanp10* graphs. For clarity we shall refer to the original unaugmented clan topology as *clanp0*.

The clan families are inspired by the many recent network formation models that begin with some regular underlying "geographical" connectivity, but then add a variable amount of random "long-distance" connectivity [13, 18]. In the clan case, we have chosen to begin with an underlying structure consisting of the four imbalanced local economies connected through the super-traders, and then stochastically layer in additional long-distance trading opportunities.

The final four topology families are closely related to more "traditional" network formation models [18]. The topologies *ERp20* and *ERp40* sample bipartite versions [9] of standard Erdos-Renyi random networks with edge probabilities of $p = 0.2$ and $p = 0.4$ respectively. Both of these models were sampled three times to obtain three different graphs.

Topologies *PA1* and *PA3* are both bipartite versions [9] of the well-studied *preferential attachment* model, in which new vertices are added incrementally and given links to the existing network, drawn according to the current degree distributions. This model is known to generate heavy-tailed (power law) degree distributions [1]. In PA1 each new vertex is given only a single initial link, thus yielding trees, while in PA3 each vertex is given three initial links, resulting in much denser networks. Again, both models were sampled three times.

Some statistics on the vertex degrees and structure of the 10 topologies are given in Figure 2, where *std* is standard deviation. The topologies pairs, cycle2 and cycle4 obviously exhibit complete symmetry across vertices, with every vertex having the same degree (1, 2 and 4 respectively). The total number of edges in these three graphs was 18, 36 and 72 respectively. For the clanp0 topology, there are 4 vertices of degree 18, 8 of degree 8, and 24 of degree 4, giving a total of 116 edges. This was relatively high, but of course clanp5 and clanp10 had even more edges. ERp40 had about as many edges as clanp5, and all other topologies had markedly fewer. Evidence of this resides in the column of average degrees.

Regarding the equilibrium properties of the 10 topologies: the pairs, cycle2, and cycle4 graphs all contain perfect matchings in every instance and thus have no equilibrium wealth variation in any instance or trial. (This is witnessed in the table by a max(equil)/min(equil) equal to 1.) The clanp0 and clanp5 topologies had no instances with perfect matchings, but all the clanp10's did. This is a systematic effect of the random extra edges present. One instance of ERp20 contains a perfect matching; the other two instances do not. The three ERp40 instances (which are denser than ERp20) all had perfect matchings.

Overall, of the 28 distinct experiments conducted, 17 used graphs with perfect matchings, which implies no equilibrium wealth variation; the remaining 11 all had varying amounts of equilibrium wealth inequality. The most imbalanced topologies were clanp0 and PA1. The table shows the standard deviation of their wealths and their range of equilibrium wealths to be the highest of all, and this inequality is a signal property. These extremes are dissipated in their denser versions.

4. SYSTEM OVERVIEW

Experiments were conducted using a distributed networked software system we have designed and built for performing a series of behavioral experiments on different games. A central server coordinates the allowed communications between players, collects data, and controls the sequencing of games.

Like most microeconomic exchange models, the model described in Section 2 does not specify an actual temporal mechanism by which trading occurs, but of course any behavioral study must choose and implement one. Before describing system and user interface details, it will be helpful to preview the mechanism implemented at a high level.

At each moment of an experiment, each subject is able to express a *limit order* that is available to *all* of the subject's neighbors; thus, like the theoretical model of Section 2, we do not permit differential pricing among neighbors (that is, prices must be on the vertices, not the edges). A limit order expresses how much of their (remaining) endowment a player is willing to trade in exchange for how much of the "other" good. Any moment at which two neighboring limit orders "cross" each other in terms of the exchange rates offered, a

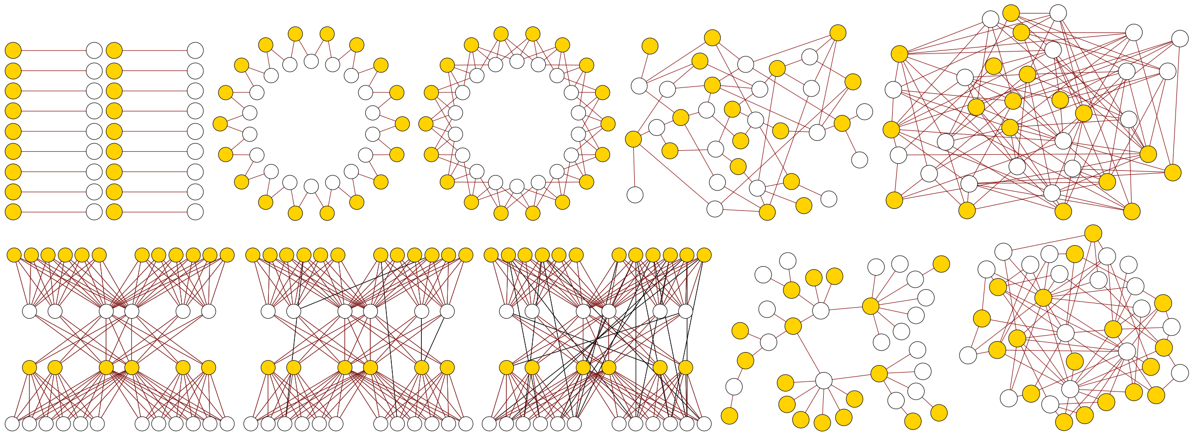

Figure 1: Sample graphs (L to R, T to B): pairs, cycle2, cycle4, ERp20, ERp40, clanp0, clanp5, clanp10, PA1, PA3

topology	*average degree*	*min degree*	*max degree*	*std degree*	$\frac{std(equil)}{avg(equil)}$	$\frac{max(equil)}{min(equil)}$	$\frac{avg(acquired)}{avg(equil)}$	$\frac{std(acquired)}{avg(acquired)}$	$L_1(equil)$
pairs	1	1	1	0	0	1	.99	.036	.016
cycle2	2	2	2	0	0	1	.87 .88 .88	.26 .29 .28	.19 .20 .21
cycle4	4	4	4	0	0	1	.96 .96 .92	.08 .12 .26	.06 .08 .16
clanp0	6.4	4	18	4.5	.72	4	.71 .77 .76	.60 .76 .66	.38 .43 .37
clanp5	6.7 7.0 6.7	4 4 4	18 18 18	4.4 4.3 4.5	.26 .13 .49	1.8 1.4 2.8	.82 .92 .91	.48 .32 .66	.27 .17 .35
clanp10	7.8 7.5 7.9	4 4 4	18 18 18	4.2 4.1 4.0	0 0 0	1 1 1	.94 .95 .92	.20 .17 .27	.12 .10 .17
ERp20	3 4.1 3.2	1 1 1	5 8 6	1.2 1.8 1.5	.22 0 .39	4 1 4	.85 .93 .90	.34 .24 .48	.21 .14 .23
ERp40	6.9 6.4 6.7	3 3 4	12 10 10	2.0 1.8 1.4	0 0 0	1 1 1	.96 .97 .93	.11 .11 .19	.06 .07 .12
PA1	1.9 1.9 1.9	1 1 1	8 8 14	1.9 1.6 2.5	1.3 .83 1.9	36 16 81	.67 .73 .77	.93 .68 1.8	.38 .40 .35
PA3	4.6 4.7 4.6	1 2 1	14 14 13	3.2 2.8 2.6	0 0 0	1 1 1	.91 .93 .87	.25 .24 .30	.17 .15 .23

Figure 2: Summary statistics of networks used. Triplets represent 3 samples of a given stochastic topology; *equil* is the vector of equilibrium wealths determined by the graph structure via the theory discussed in Section 2 and Theorem 2.1. The last 3 columns are behavioral results discussed in Section 6, not intrinsic properties of the networks.

(possibly partial) execution or trade takes place. Individuals can always see the limit orders offered by their neighbors, as well as some additional information (including the degrees of their neighbors, and the current best offers available to their neighbors). Thus each individual effectively trades against the combined book of limit orders of his neighbors. We remark that it is easy to show that the market-clearing price equilibria defined in Section 2 are indeed also Nash equilibria of the asynchronous-move, multi-party trading game that is defined by our mechanism.

All communication takes place exclusively through this trading mechanism. We now describe the user interface in detail.

4.1 User Interface

Figure 3 shows a screen shot of the game interface used by each subject in our experiments. The elements of it will be described roughly from top to bottom. The **game status** shows "pending", "in progress", or "completed", which helps the player stay synchronized as a series of different games start, stop, and change. The **elapsed time** shows the cumulative fraction of 2 minutes (the length of each individual experiment) that has elapsed since active play started. The **accumulated pay** shows the cash value of a player's winnings so far in this game.

The large section in the middle contains multiple elements. First, an **offer price** is shown as a horizontal red line at a vertical position[3] corresponding to the user's current offer price. This red line also has a number on it (in this example .69) indicating the same information.

Second, **neighbor bands** show data about individual neighboring vertices or players in the network; this example has 3 such vertical bands, indicating the presence of 3 neighboring trading partners. Each band has four sub-elements: degree, asking price, competing price, and quantity. The **degree** is a number (shown at top of these bands as "+2", "+0", and "+1") indicating the number of neighbors that vertex has *in addition to* the user. For instance, the left band indicates a neighboring vertex of degree 3, and the middle one indicates a neighbor who has no one to trade with except the user. The **asking price** appears as a horizontal line extending only the width of the band, and plotted at a vertical position that indicates the neighbor's asking price, from the perspective of this (say) Milk player as the ratio of the amount of milk demanded per unit wheat. Thus the offer price of this player and the asking prices of their neighbors are directly comparable. The asking price line also has a number sitting just above it indicating the same informa-

[3] Vertical distance was used in a nonlinear way to represent prices. By definition, the offer price for a (say) Milk player is the ratio of the amount of Milk they are offering in exchange for the amount of Wheat they are demanding, and thus lower offer prices are more aggressive. A value of 1 was plotted in the middle, while other numbers were plotted in proportion to their arctangent. This scheme gave high resolution to prices near 1, but kept extreme values from disappearing off the plot. It also represented a value and its reciprocal at equal distances above and below the middle.

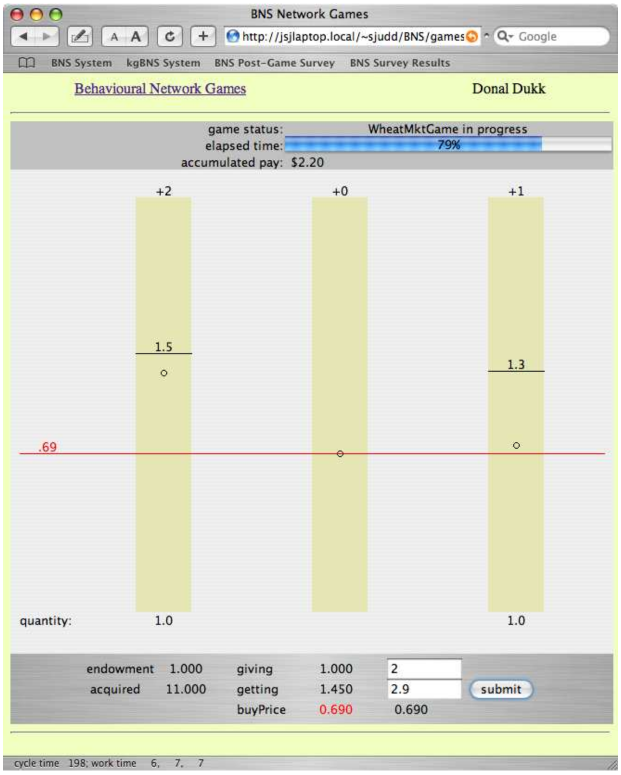

Figure 3: Screenshot of player's interface for trading.

tion. For instance, the left column has a "1.5", which means that player wishes to trade at a ratio of 1.5 units of the user's goods to 1 unit of his own. Since the user is currently offering a price of .69, there is a gap between the asking and offering prices. Visually, this is clear from the fact that the line drawn across the band is higher than the red line drawn across the whole panel. A **competing price** is shown as a small open circle, indicating by its vertical positioning the best offer currently available to this partner. The circle clearly cannot be lower than the price offered by the user, but it could be higher, indicating that one of the user's competitors is outbidding him. Also, note that the small circle cannot be higher than the asking price, or a trade would have taken place. Finally, the **quantity** is a number indicating the amount of goods being offered for sale by this partner. Thus, the full details of the standing limit order posted by the first partner is an offer to exchange 1.0 units of his goods for 1.5 of the user's. Although the ratio can be read from the price bar, the quantity cannot; the quantity field for the neighbor completes the description. Thus, subjects always know their own limit order, how it compares with those of their neighbors, what the "best price" available to each neighbor is, and how many trading partners each neighbor has.

In the lower section are 3 columns of numbers and an action button. The field called **endowment** shows the amount of remaining endowment this user still has left to trade away, and **acquired** shows the amount of the other good obtained through trades so far. The **standing limit order** appears as a column of 3 numbers called "giving", "getting", and "buyPrice". The **giving** field is the amount of the user's goods being offered for sale, **getting** is the amount of goods being requested in exchange for the quantity in the giving field, and **buyprice** is the ratio giving/getting (which is also always represented by the horizontal red line discussed above). The **new order** appears as a parallel column of 3 numbers to the right of the standing order: the top one is a field for inputting the amount of the user's goods to be offered for sale in the next limit order. Next is a field for inputting the amount of goods to be requested in exchange for the quantity in the giving field of this new limit order. Below that is the ratio giving/getting in the new limit order. Finally a **submit** button is used to post the numbers in the new order column to the marketplace so as to alter the standing order.

4.2 Game Play and Detailed Mechanism

We now provide further detail on the progression of an experiment and the execution of trades.

All players start with 10 fully divisible units of endowment, and each experiment lasted 2 minutes (120 seconds). Games are short intense affairs with all players acting simultaneously. All players have their initial standing limit order set to quantity zero. After the game starts, a player must get the cursor into the input box for a new "giving" value, and type a number in there, then (say by hitting TAB) move to the other box and enter a value for "getting". By pressing RETURN or TAB, the buyPrice field will be updated in the new order (right hand) column. The "giving" and "getting" fields could be edited at leisure again and when the player decided to commit to the new order, he would press the SUBMIT button. This immediately posts a new limit order for the player to the market server. Upon receipt at the server, the new order replaces whatever was his standing order before that. The server would deal with each newly posted order as it arrived; play was thus a series of asynchronous updates to a central server, which would relay the new state of the market to all neighboring players.

An offer to trade is transmitted to all of a vertex's neighbors simultaneously, and it is the same offer that extends to all. Again, there is no capacity for a player to offer different prices or different quantities to different neighbors. This uniformity is expressed visually by having the single red offer price line extend across all neighbor bands.

We now describe how trades are executed when an offer arrives at the server. Except for the constraint of trading only between network neighbors, what follows below largely adheres to the standard limit-order book mechanism of U.S. equities markets, in which partial executions are possible, and prices are determined by orders already standing when an arriving order causes an execution.

The player who posts the offer is momentarily considered the sole "buyer" and all his neighbors are considered "sellers". An offer arrives as a pair of numbers: giving and getting. If either of these numbers is zero then no transaction takes place. Otherwise the buyer's buyPrice is calculated as giving/getting. This price is compared against the partners' sellPrice, which is the reciprocal notion, namely their getting/giving ratios. If the incoming buyPrice is greater than or equal to the lowest partner's sellPrice then a trade takes place immediately. Some part of the buyer's endowment is transferred to that seller and some part of the seller's endowment is transferred back. The ratio of these amounts is specified by that in the seller's standing limit order. Hence the incoming order may trade for a better price than what the buyer specifies; it is always the price *previously posted* by the partner that prevails. The amounts transferred are limited by the giving limits specified in the two limit orders; neither player gives up more than his order states. One of the two players will end up with a fully cleared order, but in general the other one will have only partially cleared. If it is the buyer whose order is fully cleared then nothing else happens. If the buyer's order is not fully cleared, then a similar transaction occurs with the partner with the next lowest[4] sellPrice, and so on until either the buyer's order is exhausted or the next higher sellPrice exceeds the buyPrice. In the latter case, the remaining portion of the order is left on the market.

When the trades are completed, all the resulting changes are communicated to the affected players by altering the following GUI fields: **accumulated pay** which is always proportional to acquired, **endowment** which monotonically decreases, **acquired** which increases by how much was transferred to the player, **giving** which decreases by how much was taken from the player, and **getting** which is maintained equal to giving/buyPrice. Once traded, goods cannot be re-exchanged, re-negotiated, or re-sold.

5. HUMAN SUBJECT METHODOLOGY

We now describe our methodology in the recruitment, treatment and compensation of human subjects, which has Institutional Review Board approval at the University of Pennsylvania, and broadly follows established practices in behavioral economics [2].

36 undergraduate Penn students were recruited from a related course[5] designed and taught by author Kearns. Subjects were familiar with simple graph concepts and their role in various real-life situations, but had no prior knowledge of the particular games to be played.

A single lab with enough Linux workstations for all 36 subjects was used. Each one ran a browser in a common account that was devoid of students' personal distractions. The computer screens were arranged facing in opposite directions along long tables, so that it was difficult or impossible to see any other screen. In addition to the authors, several graduate student proctors were present during the experiments. All players were visible to proctors at all times, so any attempt to communicate via sight or sound would have been detectable. No books or electronics or any other materials were allowed anywhere but on the floor.

All players were let into the room together, and instructed to act as though they were taking an exam. No private conversation was allowed. We gave a presentation explaining the game to be played. It involved a review of the GUI, the mouse and keyboard controls, the goals of the players, the fact that graphs were generated according to different schemes not divulged, and the fact that players would be assigned to vertices in those graphs in an unbiased random fashion at the start of each experiment. We emphasized that players had no information on the global topology of any network used.

It was stressed that players' physical neighbors in the room were not necessarily neighbors in the graph, that the graph neighborhood would change with every game, and that the identities of players would not be made known during the game or at any time afterward, including all publications. Then all players logged into their machines. We did ask players to provide their name, but made it clear that the sole use of that information was to compute and distribute payments at a later date. One player's screen was temporarily projected on a large display at the front of the room while examples of play dynamics were demonstrated by asking several players to take particular limit order actions. Questions were taken and answered aloud. When all players were satisfied that they understood the purpose, mechanics, and semantics of the game and interface, we provided a sample game for them to play in which cash rewards were not given but questions were solicited.

We then started the sequence of 28 paying games. Each one was preceded by an empty screen saying "waiting for game", then the game GUI appeared on the players' screens, and we announced this aloud in the room in order to verify that everyone's machine was functioning properly. Players became familiar with the local structure of their neighborhoods in the upcoming game. Their interfaces were live and would take inputs, but the server would not yet process orders. After a few seconds, we announced the beginning of the game and pressed a key on the server to enable play. A small bell rang on each computer, followed by a silent 2 minute period, whose progress was displayed on the elapsed time display at the top of every players' applet GUI. A small bell rang when time was up, and we also announced the end of play aloud. The screens remained frozen in their final game state so that players could take note of it. After a moment, at a command from the server console, the screens reverted to the "waiting" display. On average, about a minute was spent waiting for the next game to be ready.

After all the training preamble, we conducted 28 paying games in 85 minutes. We paid a cash incentive equal to US$2 per 10 units of the neighbors' goods that the player was able to trade for. Thus, for example, a player who managed to trade their initial endowment for 15 units of the other good was paid $3, and one who managed to trade for 7.5 units was paid $1.50, and so on. Since each of the 36 players started with 10 units of goods, the maximum social welfare (total payout) was $72 during each 2 minute game. After all 28 games were completed, we asked players to fill out an online survey form to record observations, strategies, complaints, and suggestions about the games, the preparation, the equipment, and the general event. On average, players earned about $50 in 2 hours.

6. RESULTS AND FINDINGS

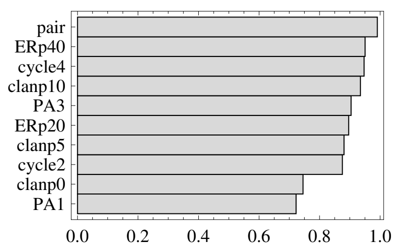

Figure 4: Mean fraction of maximum wealth realized by topology.

Collective Performance and Topology. We begin our findings with a discussion of absolute overall performance

[4] Ties amongst sellers' sellPrices are resolved by taking the one with the smaller vertex index first. This results in a small bias which was noticeable to an omniscient observer, but was imperceptible to the players. Its effect could always be overcome by raising one's prices by a miniscule amount. We believe the bias was inconsequential to our results.

[5] See www.cis.upenn.edu/˜mkearns/teaching/NetworkedLife

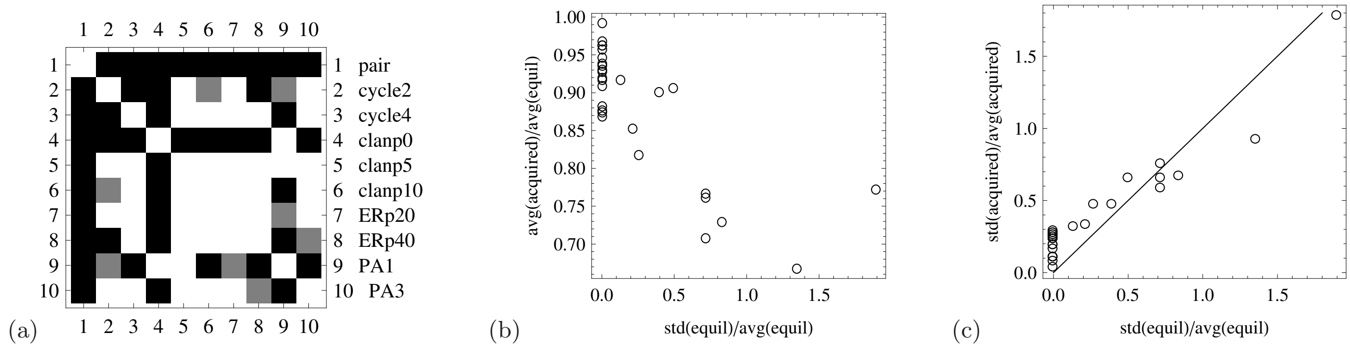

Figure 5: (a) t-tests between mean wealths by topology. (b) equilibrium wealth standard deviation vs. empirical wealth realized. (c) equilibrium wealth standard deviation vs. empirical wealth standard deviation. Note the clusters of points at $x = 0$, corresponding to the networks containing perfect matchings. The data for plots (b) and (c) are in the table of Figure 2.

and its relationship to network topology. In Figure 4 we show a bar plot of the fraction of the maximum possible social welfare (\$72) empirically earned by topology type, averaged over the three trials for each topology (1 for the pairs network). Overall performance was strong, with an average social welfare over all 28 experiments of 0.88 of the maximum possible; the maximum for any individual experiment was 0.99 (pairs network) and the minimum was 0.67 (for a PA1 network)[6]. We thus conclude that networks of human subjects can indeed come close to clearing the market using only local information and incentives.

Turning to the effects of topology, Figure 4 shows large differences in market-clearing performance between topologies, from an average of only 0.72 for the PA1 trials to the 0.99 for the pairs network. While three trials per topology is too few to establish statistical significance, remember that each trial itself averages the 36 individual wealths. For each topology, we can compute the $3 \times 36 = 108$ (36 for the pairs topology) individual wealths realized, and compare these sets across pairs of topologies in a t-test for different means[7]. Figure 5(a) shows in black those pairs with unequal means at statistical significance $p \leq 0.05$, and in gray those pairs at $p \leq 0.1$. The overall conclusion is that many topology pairs can be distinguished by this test, and thus topology does indeed influence collective performance. Looking in slightly more detail, we see that the pairs topology is strongly distinguished from all others, while clanp0 is distinguished from all but PA1; there is also a group (clanp5, clanp10, ERp20, ERp40) none of which can be distinguished from each other.

Equilibrium Predictions, Collective Level. We know from Section 2 that topology strongly influences theoretical equilibrium wealth distribution. Does this influence have systematic behavioral effects, and if so, what are they?

Figures 5(b) and 5(c) together show some of our most striking findings. The first scatterplot represents each of the 28 experiments as an open circle. The x coordinate measures the standard deviation of the equilibrium wealths (expressed as multiples of the initial endowments) — an entirely theoretical quantity that we view as a measure of the wealth imbalance required to clear the market at equilibrium. It varies from 0 (for all those networks containing perfect matchings) to 1.89 (almost twice the endowment). The y coordinate shows the actual empirical fraction of the maximum possible wealth realized. Figure 5(b) demonstrates that these two quantities are highly negatively correlated — their correlation coefficient is -0.8 with a p-value below 0.001. Even after omitting the 17 networks with perfect matchings clustered at $x = 0$, the negative correlation remains strong (-0.64 with $p \approx 0.03$). Thus, we see that *collective behavioral performance degrades as the theoretical equilibrium predicts greater wealth disparity.*

Figure 5(c) has the same x coordinate, but now the y coordinate plots not the behavioral social welfare, but the standard deviation of the empirical wealths (expressed as multiples of the average wealth obtained). The diagonal line represents identical standard deviations in the equilibrium and empirical wealths. Here we see a strong *positive* correlation — greater equilibrium wealth imbalance does indeed correctly predict greater empirical wealth imbalance. The correlation coefficient is a whopping 0.96, again with a p-value below 0.001; it remains 0.95 with a similar p-value even after omitting networks with a perfect matching. Thus, despite greater equilibrium inequality causing lower realized wealth, it does indeed seem to cause greater empirical inequality over whatever wealth is earned.

Of course, this strong correlation of the equilibrium and empirical standard deviation does not yet imply that the empirical wealth is imbalanced in the same *way* (that is, across the same individuals) as the equilibrium. That question is examined next.

Equilibrium Predictions, Individual Level. So far we have seen that the equilibrium theory can predict some macroscopic aspects of the collective behavior; how well does it do at individual-level modeling?

One basic approach to answering this question is to compute the L_1 (variation) distances between the normalized empirical wealth distribution and the normalized equilibrium wealth distribution for each of the 28 experiments, and to compare these distances with those yielded by other plausible predictors of the behavioral outcomes. As a baseline, we note that the mean L_1 distance of random (normalized) vectors to the 28 empirical wealth distributions is approximately 0.59. In contrast, the distances for the equilibrium predictions are considerably smaller[8], with a mean of just

[6] Values for all individual experiments can be found in the *avg(acquired)/avg(equil)* column of Figure 2.

[7] All t-tests reported are unpaired, two-tailed and unequal variance, the most conservative possible.

[8] Values for all individual experiments can be found in the $L_1(equil)$ column of

0.22 .

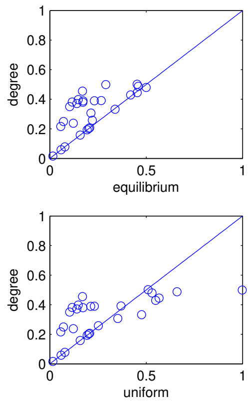

Figure 6: Comparison of models of wealth distribution.

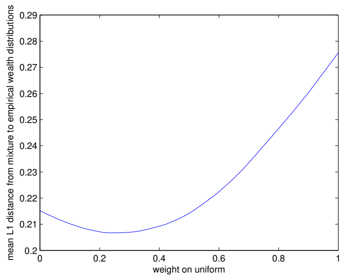

Figure 7: L_1 mean of mixture model vs. α.

Furthermore, by this measure the equilibrium predictions outperform other plausible theories. We first compare the L_1 distances from empirical wealths to equilibrium predictions against the distances to the degree distribution. The degree distribution makes the reasonable prediction that greater wealth will simply follow those with more trading opportunities. In the upper-left scatterplot of Figure 6 are 28 plotted points, one for each experiment. The x value is the L_1 distance between the equilibrium distribution to the empirical distribution for that experiment, while the y value is the L_1 distance between the degree distribution and the empirical distribution. Points above the diagonal are a victory for the equilibrium model, and we see that the equilibrium L_1 distances are equal to or smaller than those of the degree distribution in *all 28 experiments*. The mean L_1 for the degree distributions is 0.32, which is distinguished in a t-test from that of equilibrium (0.22) at significance $p < 0.0001$.

Given the suggestion of inequality aversion made by Figure 5(b) above and discussed further below, it is also reasonable to simply predict that any wealth earned will be distributed uniformly in each experiment, which yields an average L_1 of 0.28 — actually better than degree distribution (though not statistically significant; lower-left scatterplot of Figure 6), but again the same or worse than equilibrium predictions in all 28 cases (upper-right scatterplot of Figure 6). The equilibrium mean L_1 passes a t-test for being smaller than that of the uniform distribution at $p \approx 0.03$.

Finally, foreshadowing our finding below that one of the main behavioral deviations from equilibrium theory is an excess of 1-for-1 trading, we find that our best model to date is one of the form $(1-\alpha)\vec{e} + \alpha\vec{u}$, where $\vec{e}$ is the equilibrium wealth distribution for the network under consideration and $\vec{u}$ is the uniform distribution. We find that the best mean L_1 occurs for $\alpha \approx 0.25$ and yields an average L_1 of 0.21 — better than equilibrium alone, but not statistically significant (lower-right scatterplot of Figure 6). Figure 7 shows the performance vs. α, which exhibits a distinct minimum.

Behavioral Dynamics. Thus far we have focused on the final outcomes or wealths of an entire experiment, at both the collective and individual levels. But how does the action typically unfold *within* a particular game? We now turn to such questions of behavioral network dynamics.

We first examine the raw level of activity as a function of time within an experiment. The first plot in Figure 8 shows the average, over all 28 experiments, of the number of changes to their outstanding bids made by players in each 10-second block of an experiment. Note that such changes may or may not result in any actual transactions taking place. The most striking feature of this plot is the large volume of bid adjustments being made in the first 10 seconds or so — by far the largest of any period. However, there remains a large amount of activity distributed relatively evenly through the remainder of the experiment.

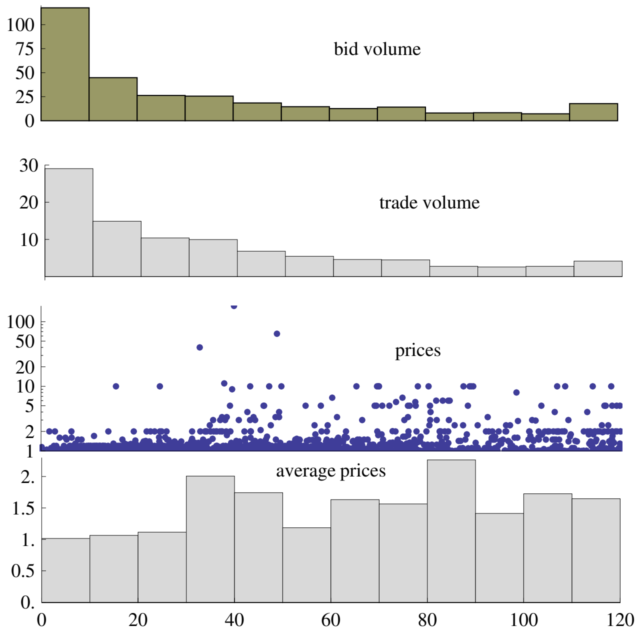

Figure 8: Evolution of activity summarized for all games. Horizontal axis is time, in seconds.

The second plot in Figure 8 shows the average number of *executed* trades that took place in each 10-second interval; there are naturally fewer than the number of bid adjustments, but again there is an early rush of executed trades, followed by reduced but still significant trading. Interestingly, neither the bid adjustments nor the executed trades plots show any evidence of a late "rush" to execute trades as time expires.

Further insight into the initial burst of trading is gleaned from the third and fourth plots in Figure 8, which represent trading prices rather than just volume. The third plot represents each executed trade by a single point whose x value

Figure 2.

is the time of execution and whose y value is the price enjoyed by the "winner" in that trade — that is, if a units of wheat were traded for b units of wheat, the plotted value is $\max(a/b, b/a) \geq 1$. Note the log-scale y axis. The plot aggregates all trades over all 28 experiments; the fourth plot shows averages of these values within each 10-second interval. It is clear that the early rush of trading takes place at highly equitable exchange rates (≈ 1), while later trading tends to contain a sizable fraction of inequitable trades ($>> 1$). Indeed, the average winning price in the first 30 seconds of play across all experiments was 1.05, while the average over the final 90 seconds of play was 1.71; this difference is highly statistically significant. This property of early trading may (partially) explain our earlier finding that the best model for final wealths tilts the equilibrium theory towards equality.

There are at least two plausible explanations for the preponderance of early, equitable trades. It might be that some players jump at equitable trades quickly to avoid the risk of being "locked out" and unable to trade their endowments later. Alternatively, it may just take more time for players to be convinced of the need to demand and pay higher prices, since in the equilibrium theory this is distinctly a network effect that may thus require some time to propagate.

We note in passing that there appears to be a psychological barrier price of 10:1, which is visible in the line of 20 dots at 10 in the middle plot. Only 4 trades out of a total of 2746 in 28 games were above that price.

Fragmentation of Liquidity. One of the most interesting behavioral phenomena we observed was the isolation or fragmentation of goods due to structural "traumas" to the network induced by trading. To describe this we need to introduce the notion of Conditional Equilibrium Wealth (CEW).

The maximum social welfare achievable and the equilibrium prices change over time due to trading, and it is instructive to see an example of how this can happen. In the case of a 3-vertex line graph with endowments of $\{10, 10, 10\}$ where the middle vertex is able to trade with both end vertices, the equilibrium price is 2:1, the equilibrium wealths are $\{5, 20, 5\}$, and the maximum social welfare is 30. If a 5-for-5 trade takes place on the left link, the new *conditional* equilibrium wealths change to $\{5+\frac{1}{3}5, 20, \frac{2}{3}5\}$, and the maximum social welfare is unchanged. If a 10-for-10 trade takes place, though, the new conditional equilibrium wealths are $\{10, 10, 0\}$, and the maximum social welfare drops to 20 — the 10-for-10 trade effectively isolates the right-hand vertex, who has no trading partner remaining.

Hence, just as equilibrium calculations are pertinent *a priori* predictions of the final wealths of players at the beginning of a game, they are also valuable for examining the evolution of a game. More formally, we define the CEW at time t for each player to be the sum of its acquisitions up to time t and its new equilibrium wealth given the remaining endowments of all players at time t. We show empirical plots of these 36 time-evolving values in Figure 9 for two of our experiments, an ERp40 trial and a clanp0 trial. The starting value of each player's curve is by definition their original equilibrium wealth, the ending value is what that player actually acquired in the experiment, and in between their CEW can go both up and down. The uppermost curves plot the summed CEW values over all players, which we discuss shortly.

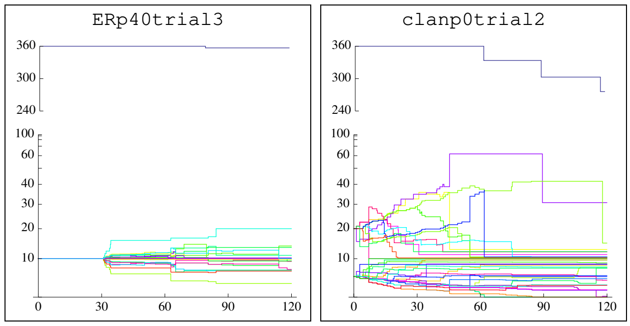

Figure 9: Summed CEW values (upper plots) and individual values (lower) evolving over time, for two games. Horizontal axes are in seconds.

It is clear from these two examples that rather different dynamics are in force in the two experiments, especially with regards to the volatility of the CEWs. More generally, we find that the standard deviations of the CEW curves on a per-player basis are strongly influenced by the topology of the network. Given a set of standard deviations for each player's CEW data in each of our 10 different network topologies, a t-test will find almost every pair of them significantly different at the $p = 0.05$ level.

The *sum* of all players' CEW values is the maximum social welfare that is still possible at that point in time. As long as the market can still clear this will remain a constant, but whenever goods become isolated from their potential buyers, it falls. Such (irreparable) fragmentation to the network causes the sum to drop in sudden steps. Figure 10 presents all the summed CEW traces for all experiments, grouped by topology. Note that the different topologies produce dif-

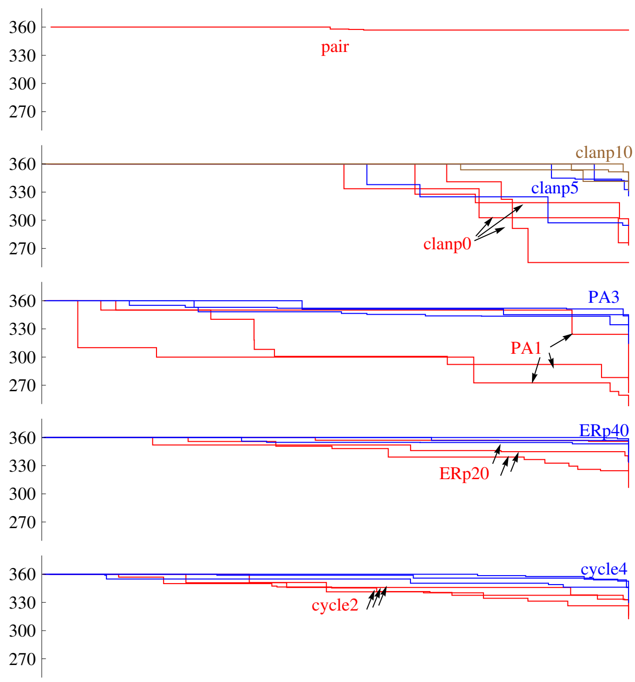

Figure 10: Summed CEW values for all experiments by topology. Horizontal axis is from 0 to 120 secs.

ferent amounts of trauma. The clan and PA graphs show more than the others. The PA1 graphs are only 1-connected and show damage early, whereas the clan graphs are all 4-connected and damage does not occur until much later. In general, adding edges to a graph delays and reduces damage. (Edge count relations are thus: clanp0 $<$ clanp5 $<$ clanp10; PA1 $<$ PA3; ERp20 $<$ ERp40; and cycle2 $<$ cycle4.) Note also that clanp0 has many more edges than cycle4, and in spite of this it shows much more damage; structure matters.

Regret in the Final Outcomes. We conclude with one more analysis of the final wealths of the players. All games ended with some players having unsold endowment. We use the term "aggressive regret" to mean the amount of endowment that a player held at the end of a game (since they were unable to trade it at the rates they were demanding), and the term "passive regret" to mean the total amount of endowment remaining in all of a player's neighbors (since this is what in principle was available to them right before time expired). It occasionally happened that a player would have both types of regret, which we call "chicken" regret (since trades were possible but players chose to not engage in them before time expired).

Figure 11 shows regret values for all 9×36 players of all 9 clan experiments. Each player is represented as a ball with a vertical bar running from it to the zero line. Passive regret is shown as a negative value, aggressive regret as positive. When a player has both types of regret, they are shown colinearly. The players have been sorted first by their starting equilibrium wealth; in the case of ties, they are further sorted by passive regret; in the case of further ties, they are sorted by aggressive regret.

There was a marked tendency for players with a high starting equilibrium wealth to leave neighbors with surplus goods, and for players with a low starting equilibrium wealth to fail to sell all their goods. An alternative phrasing is that many disadvantaged players were unwilling to pay very high prices, and many advantaged players were unwilling to charge very high prices. This phenomenon was most pronounced in the clanp0 networks, and is again broadly consistent with the collective inequality aversion we have discussed.

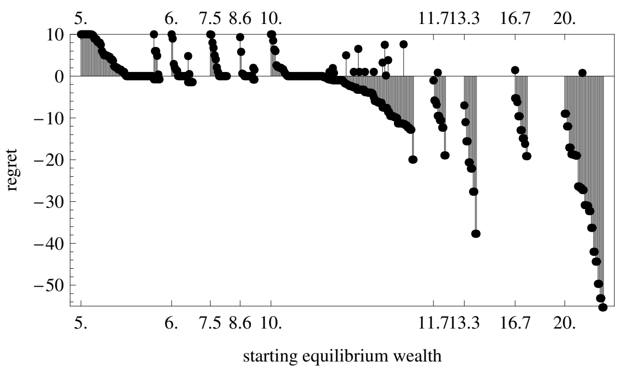

Figure 11: Regret in clan graphs. Values on the horizontal axis show equilibrium wealth, but successive players with equal equilibrium wealth are shown offset by a tiny amount to the right.

Regret is intimately related to social welfare and structural trauma. The sum of all the aggressive regret and the final social welfare is equal to the original maximum social welfare (360 units). With the exception of the small amount of chicken regret, though, all the bars in Figure 11 are the result of structural damage to the network through trade. These are not cases of players sitting on their hands and waiting for lower prices; there is simply nothing they can do near the end of the game to alleviate their excess supply or demand.

7. REFERENCES

[1] A. Barabasi and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

[2] C. Camerer. *Behavioral Game Theory.* Princeton University Press, 2003.

[3] G. Charness, M. Corominas-Bosch, and G. Frechette. Bargaining and network structure: An experiment. *Journal of Economic Theory*, 136, 2007.

[4] N. R. Devanur, C. H. Papadimitriou, A. Saberi, and V. V. Vazirani. Market equilibrium via a primal-dual-type algorithm. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations o f Computer Science*, 2002.

[5] D. Gale and S. Kariv. Trading in networks: A normal form game experiment. 2007. Working paper.

[6] M. Jackson. The economics of social networks. In *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society.* Cambridge University Press, 2005.

[7] S. Kakade, M. Kearns, J. Langford, and L. Ortiz. Correlated equilibria in graphical games. In *Proceedings of the 4th ACM Conference on Electronic Commerce*, pages 42–47. ACM Press, 2003.

[8] S. Kakade, M. Kearns, and L. Ortiz. Graphical economics. In *Proceedings of the 17th Annual Conference on Learning Theory*, pages 17–32. Springer Berlin, 2004.

[9] S. Kakade, M. Kearns, L. Ortiz, R. Pemantle, and S. Suri. Economic properties of social networks. In *Proceedings of the Eighteenth Conference on Neural Information Processing Systems*, 2004.

[10] M. Kearns. Graphical games. In N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani, editors, *Algorithmic Game Theory.* Cambridge University Press, 2007.

[11] M. Kearns and S. Suri. Networks preserving evolutionary equilibria and the power of randomization. In *Proceedings of the 7th ACM Conference on Electronic Commerce*, pages 200–207. ACM Press, 2006.

[12] M. Kearns, S. Suri, and N. Montfort. A behavioral study of the coloring problem on human subject networks. *Science*, 313(5788):824–827, 2006.

[13] J. Kleinberg. Navigation in a small world. *Nature*, 406:845, 2000.

[14] R. Kranton and D. Minehart. A theory of buyer seller newtorks. *American Economic Review*, 91(3):485–508, 2001.

[15] C. Papadimitriou. The complexity of finding Nash equilibria. In N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani, editors, *Algorithmic Game Theory.* Cambridge University Press, 2007.

[16] T. Roughgarden. *Selfish Routing and the Price of Anarchy.* MIT Press, 2005.

[17] E. Tardos and T. Wexler. Network formation games and the potential function method. In N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani, editors, *Algorithmic Game Theory.* Cambridge University Press, 2007.

[18] D. J. Watts. *Six Degrees: The Science of a Connected Age.* W. W. Norton, Cambridge, Mass., 2003.